DOCUMENT RESUME

ED 089 865 PS 007 220

AUTHOR Quarton, Carol Anne
TITLE The Development of a Test of Comparative Language for Young Children.
INSTITUTION Pennsylvania State Univ., University Park. Inst. for the Study of Human Development.
SPONS AGENCY Pennsylvania State Dept. of Public Welfare, Harrisburg.
PUB DATE 1 Jun 72
NOTE 71p.; The Pennsylvania Day Care Study Project, Technical Report #4; For other documents in this series, see PS 007 218 through PS 007 224

EDRS PRICE MF-$0.75 HC-$3.15 PLUS POSTAGE
DESCRIPTORS *Behavioral Science Research; *Evaluation; *Language Development; Language Skills; *Measurement Instruments; *Preschool Children; Tests
IDENTIFIERS *Pennsylvania Day Care Study Project

ABSTRACT

The purpose of the study was to begin the development of a test of comparative language which would have potential usefulness as an evaluation instrument for early childhood education programs. The study provides preliminary information concerning reliability, validity, and practicality of two parallel forms of the test. Included in each form are items concerning position, color, shape, quantity, size, and compound expressions of relational terminology requiring simultaneous use of two dimensions. The two forms were found to be highly reliable over time and internally consistent. Preliminary validity information uncovered appropriate relationships with age, socioeconomic status, and mental maturity. This study was conducted as part of the Pennsylvania Day Care Study Project. (Author/DP)

U S DEPARTMENT OF HEALTH
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION
THIS DOCUMENT HAS BEEN REPRO
DUCED EXACTLY AS RECEIVED FROM
THE PERSON OR ORGANIZATION ORIGIN
ATING IT. POINTS OF VIEW OR OPINIONS
STATED DO NOT NECESSARILY REPRE
SENT OFFICIAL NATIONAL INSTITUTE OF
EDUCATION POSITION OR POLICY

ED 089865

THE PENNSYLVANIA STATE UNIVERSITY

The Institute for the Study of Human Development
Center for Human Services Development

THE DEVELOPMENT OF A TEST OF COMPARATIVE LANGUAGE FOR YOUNG CHILDREN

PS 007220

Prepared by:

Carol Anne Quarton

THE PENNSYLVANIA DAY CARE STUDY PROJECT
Technical Report #4 June 1, 1972

ED 089865

THE DEVELOPMENT OF A TEST OF COMPARATIVE LANGUAGE FOR YOUNG CHILDREN

Carol Anne Quarton

The Pennsylvania State University

June, 1972

This report was partially supported under contract with the Pennsylvania Department of Public Welfare. The opinions and recommendations expressed are those of the authors and do not necessarily reflect the official position of the sponsoring agency.

ABSTRACT

The purpose of the study was to begin the development of a test of comparative language which would have potential usefulness as an evaluation instrument for early education programs. The study provides preliminary information concerning the reliability, validity and practicality of two parallel forms of the test. Included in each form are items concerning position, color, shape, quantity, size and compound expressions of relational terminology requiring simultaneous use of two dimensions. The two forms were found to be highly reliable over time and internally consistent. Preliminary validity information uncovered appropriate relationships with age, socio-economic status and mental maturity.

CONTENTS

		Page
I.	INTRODUCTION	1
II.	A REVIEW OF THE LITERATURE AND THE PROBLEM STATEMENT	4
	A Review of the Literature	4
	The Problem Statement	12
III.	PRELIMINARY STUDY	14
	Purpose	14
	Method	14
	Results	16
	Discussion	24
IV.	THE MAIN STUDY	26
	Purpose	26
	Method	26
	Results	28
	Discussion	53
V.	SUMMARY	56
	Characteristics of the Test	56
	Hierarchy of Comparative Language Development	57
REFERENCES		59
APPENDIX		62

CHAPTER I

INTRODUCTION

The purpose of this study was to develop a language test that would be useful in assessing the development of a young child's use of comparatives. Such an instrument has potential as both a pretreatment diagnostic device and as a posttreatment evaluation mechanism. It may also provide information leading to a greater understanding of the hierarchical development of language skills.

Language has become an area of intense interest and importance to early childhood educators over the past 15 years (Frost, 1968; Baratz, 1968; Williams, 1970). Although it has traditionally been a vehicle of educational instruction, this resurgent interest in its potential importance as a program component of instruction has caused early childhood educators to cite the need for truly effective language components within educational programs for the young (Williams, 1970; Bailey, 1966; Blank, 1970; Englemann, 1970).

Language programs for young children do exist. They vary, however, as to what is considered important in terms of content and methodology. Some language programs, such as the one in Bereiter and Englemann's "Distar" (1969), contain very specific goals, rationales, and implementation procedures. Others are so ill-focused that it is difficult to see how they can function as a viable part of a total program.

The vagueness of many language programs may be partially

the result of the controversy among educators concerning the role language plays in intellectual development. Many educators feel that language development is crucial in facilitating cognitive growth through verbal mediation (Bernstein, 1968; Hess and Shipman, 1965). Others feel that language is an overt sign of intellectual competency and growth rather than a prime causal factor (Frost and Rowland, 1968; Piaget, 1926). Although there is ambiguity concerning the exact relationship between thought and language, there is general agreement that a close relationship between the two exists.

Many programs for young children currently being prepared include concept goals associated with categorization, perception, seriation, and simple logic (Lavatelli, 1970; Kamii, 1970; Nimnicht, 1969). The language associated with such content learning has, however, only been peripherally treated as a viable curriculum component in itself, despite growing evidence that such language may be a prerequisite for such concept formation and extension (Inhelder, Bovet, Sinclair, and Smock, 1964; Sinclair-de-Zwart, 1969; Griffiths, Shantz and Sigel, 1967).

One aspect of language that is thought to be important to general intelligence and concept learning is comparative language. Words denoting color, shape, size, position, and quantity, utilized alone and in combination, are useful in applied content areas like categorization, perceptual discrimination, reading readiness, number readiness, conservation, and seriation skills. However, utilizing

comparative language in an early childhood education program presents a problem. The information and abilities an individual child brings to school need to be determined. Once the initial data are gathered, intermittent monitoring of the child's progress might be useful in developing productive alternatives that would stimulate comparative language competencies.

A comparative language test is one way of meeting the needs of comparative language utilization. The test could provide a pre-, inter-, and posttreatment evaluation device. In addition, it could supply data which would aid in determining the developmental sequence of comparative language concepts. This sequence might subsequently serve as a model for devising systematic language program components.

CHAPTER II

A REVIEW OF THE LITERATURE AND THE PROBLEM STATEMENT

A Review of the Literature

Language development has long been viewed as a significant factor in human development. Consequently, the topic has been examined in many different contexts. Developmental psycholinguists, for instance, have studied models of language and language acquisition in order to explain the manner in which language is perceived and produced (Menyuk, 1970). Current thinking in the discipline centers on the nativist view of language acquisition, which stresses that children are biologically predisposed to develop language and that environment triggers rather than produces stages of development. Language acquisition is seen as a process by which linguistic knowledge and abilities develop in the child. It is asserted that development can best be described by discovering the types of linguistic rules operating in the child's language repertoire and by determining the relationship between them and maturational factors (Williams, 1970).

Language development has also been studied from a sociological viewpoint. Sociolinguists have regarded language as a cultural phenomenon. As a cultural force which serves as a means of transmitting social structures, language is crucial in the socialization process. Yet conversely, language is also partially a product of socialization. It is associated with a number of

variables, including social status, ethnicity, and other circumstances which influence the speech situation. Williams (1970) cites an emerging generalization in sociolinguistic theory specifying that normal development of a child's language must be viewed relative to the demands of his primary speech community. In other words, an individual evolves within a setting. The child and his environment interact and shape each other. Demands made on the child within the context of his immediate environment determine modes of communication optimally productive for him.

None of the current theories of language development adequately explain language acquisition in relation to innate ability, biological maturation, motivation, and cultural-environmental factors. Neither is the relationship between language and cognition adquately explained. This study cannot resolve these issues. It can, however, add to the description of the acquisition of comparative language in young children.

PS 007220

Comparative language includes words and phrases which categorize and/or describe objects within specific contexts. Some examples of comparative language are represented by the following list: open, bigger, red, tallest, rectangle, beside, many. They are verbal symbols representing simple concepts which almost every child must learn in conjunction with perceiving and interpreting his world.

Some forms of comparative language are learned by almost all individuals regardless of culture, and it is assumed that the

sooner in life these elements are learned and internalized, the sooner more complex concepts can evolve (Palmer, 1968). Bernstein (1968) and Hess and Shipman (1965) suggest that there is a definite advantage in being able to utilize an elaborate code which contains specific words to describe objects and situations.

Many studies (Bruner, 1964; Beilin, 1965; Wohlwill, 1960) suggest that a child must have an understanding of comparative language concepts in order to deal more effectively with tasks in which comparative relationships are an integral part. This necessity of comparative language competency is particularly evident in many Piagetian tasks which require a verbal justification as part of the criterion for concept mastery. Berko and Brown (1960) consider Piaget's developmental sequence to reflect vocabulary growth almost exclusively. Braine (1959, 1964) suggests that Piaget's tasks sometimes demand a terminology grasp which actually exceeds the conceptual grasp the task is designed to measure.

Sinclair-de-Zwart (1969) investigated the relationship between comparative language and the development of the first concrete operations theorized by Piaget. Her objectives were twofold: 1) to see if modifications in a child's thinking process as a result of the first concrete operations are paralleled by linguistic development, and 2) if a child's thinking is paralleled by linguistic acquisition, to see if a child who lacks a certain concept or operation would show operatory progress after having undergone verbal training.

The results of the first part of Sinclair-de-Zwart's study were consistent with the theory that conservation of discrete units is acquired before that of continuous quantities. After pretesting the subjects, she divided them into three groups: conservers, intermediate conservers, and nonconservers. Results indicated that children with conservation used comparatives 90-100 percent of the time in describing materials, while children without conservation used absolute terms (e.g., a lot, a little) 90 percent of the time. Twenty percent of the nonconservers used comparatives when describing descrete units, but they did not use them for continuous quantities.

Of the children who did not conserve, 75 percent used the same term to describe two different dimensions (e.g., using "fat" to describe both "long" and "thick" variables). Eighty percent of the children who conserved used differentiated comparative terms to coordinate the two dimensions of the materials.

In the second part of the experiment Sinclair-de-Zwart attempted to teach the nonconservers the comparative expressions utilized by the conservers and tried to determine what effect the training had on their operational level. She found it easy to teach the children differentiated comparatives (e.g., longer, thin, short), difficult to teach "more" and "less," and still more difficult to teach the children to utilize two coordinated comparatives (e.g., long and thin). Although only 10 percent of the children exhibited operational progress, it appeared that the children noticed and

verbalized the appropriate covarying dimensions involved in the task.

Sinclair-de-Zwart concluded that: 1) language acquisition does parallel operational structuring, and 2) verbal training leads subjects without conservation to direct themselves to the pertinent aspects of the problem.

The results of the study by Sinclair-de-Zwart tend to agree with those of Inhelder, et al. (1964). Training in the use of comparatives operates to direct the child's interactions with the environment and thus draws his attention to the relevant dimensions of task situations. They go on to comment that observed changes in the justifications given for an answer in conservation tasks indicate that language does aid in the storage and retrieval of relevant information.

Gruen (1966) concludes that verbal pretraining on "more," "longer," and "some" had the effect of screening out perceptual cues leading to a greater percentage of five-year-old children making conserving responses. Halasa (1969) contends that, for a sample of disadvantaged children, an understanding of the concepts "more," "less," and "same" may be a necessary but not sufficient condition for number conservation.

Evans (1971), after a review of relevant literature, asserts that, if the classical Piagetian verbal testing techniques are used, an investigator should first determine the subjects's ability to discern similarities and make appropriate use of relevant vocabulary.

Otherwise ". . . results could be an artifact of questioning. In short, it would become difficult to determine whether: 1) a child cannot conserve, 2) a child can conserve but is unable to decipher a question, or 3) he can neither conserve nor decipher a question tion" (p. 217).

Griffiths, Shantz and Sigel, (1967) examined the ability of preschool children to use the terms "more," "less," and "same" when comparing the number, length, and weight of objects. "Same" was found to be used significantly less often than "more" or "less." The correct use of "same" was highly dependent upon the content area, while "more" and "less" were used more frequently across content areas. The frequency of correct "same" responses was greatest for length, next came number, and last was weight. This order parallels the order of difficulty of the concepts. Sinclair-de-Zwart's results were similar concerning continuous and discrete quantities indicating parallels between the acquisition of language of operational structures.

This information indicates that a language development difficulty sequence may exist which parallels concept difficulty hierarchies. Psycholinguistic theory, as discussed by Williams (1970) and Menyuk (1970), certainly suggests that language acquisition may be hierarchical. As language concepts are acquired, they are stored for further use and serve as part of a growing foundation upon which more complex language components are integrated into the language system.

The studies previously cited make it evident that comparative language is an important aspect of general language acquisition serving as an aid in determining specific cognitive growth. However, relatively few comparative language measures exist. Bussis (1968) developed a comparative language test for the Educational Testing Service. It was later adapted by Peters (1968) for his Language Comprehension Task. Peters retained the basic design, the standard procedures for administration, and the original 20 items plus two warm-up items; he added 10 new items but used the same format. Content of the test included present, past and future verb tenses, negatives, same and different, and words denoting quantity, time and size. Peters utilized the test as a pre- and posttreatment evaluation device.

Montemerlo (1970) drew upon Peters' work and developed two tests. One, the Test of Usage of Number Conservation Terminology, consisted of 10 language concepts: same number, different number, more, fewer, long, short, big, small, close together, far apart. The other, the Test of Usage of Distance Terminology, included the following 10 language concepts: near, far, fast, slow, lot of time, little bit of time, big, small, back and front. Both tests consisted of three parts; each utilized the 10 language concepts included in the test. The first part measured the ability of the subject to use each of the 10 number-related terms in the absence of a perceptually misleading cue. The second measured the ability to use the 10 number-related terms in the presence of a perceptually

misleading cue. The third part measured the subject's ability to coordinate pairs of these terms to solve problems. Montemerlo utilized the tests as pre- and posttreatment evaluation mechanisms.

Palmer (1968) devised the Concepts Familiarity Index to measure knowledge of those concepts incorporated into an early childhood intervention program. The following nine dimensions were included in the curriculum: position, direction, sensory-tactile, quantity, movement, size, form, same, and different. The pretest scores were used to determine a sequence for instruction based on the relative difficulty of the items. The test was also used as a posttest evaluation device, although the test did not retain its original form.

None of the comparative language tests cited are readily available in an easily useable form. They were developed to meet specific research needs in the absence of suitable available tests. Because they were built to fit prescribed needs, they do not have the scope and depth to tap general comparative language proficiency.

Summary. This brief review of the literature indicates that a relationship does exist between comparative language and conceptual problem-solving tasks. Children who have grasped the rudiments of comparative language seem better able to focus on the relevant dimensions of a task, coordinate two discrete comparative dimensions, and produce conserving responses.

There is also some indication that a comparative language hierarchy may exist.

In addition, it seems that there is a need to develop a valid and reliable index of comparative language development that would have the scope to meet a variety of needs in research, assessment, and curriculum planning purposes.

The Problem Statement

The purpose of this study was to develop a test of comparative language which would: 1) meet the standard psychometric requirements of reliability and validity, 2) be practical as an assessment instrument with young children ages three to five, and 3) have potential value in program evaluation and curriculum development. Consideration was given to time requirements of testing, utilization ease, format, and content. Two parallel versions were developed.

Additionally, evidence was sought which would prove helpful in defining the hierarchy of the acquisition of comparative language.

Further, for this particular study, formulating hypotheses is relevant only when considering validity estimates for the newly developed measuring instrument. Since the instruments developed reflect the developmental acquisition of language (a specific kind of language in this case), it could be hypothesized that:

1. Within a sample of three to five-year-olds, older children will attain a higher overall score on the comparative language measure than younger children will.
2. Within the sample there will also be a significant correlation of the overall score on the comparative language measure with IQ or mental age.
3. Since the literature indicates the context in which language is learned is related to the type of language learned and since middle-class children learn more elaborated language codes, middle-class children will score significantly higher than lower-class children.

Finally, indications of a hierarchy of development for comparative terms were sought.

CHAPTER III

PRELIMINARY STUDY

Purpose

Following the work of Bussis (1968), Peters (1968) and Montemerlo (1970), a preliminary form for a comparative language test was developed. This version of the test consisted of 125 items within seven categories: color, shape, size, quantity, position, color and shape, and compound relational (a category comprised of paired comparatives of size, quantity and position and same and different). The purpose of the study was to obtain preliminary data concerning the reliability, practicality, and effectiveness of the item content and item form.

Method

Sample. Twenty children enrolled in one early childhood education program at the Pennsylvania State University served as subjects for the study. The sample contained an equal number of males and females. The subjects ranged in age from three years, ten months, to four years, eight months. The mean age was four years, one month. The children were all white and predominantly middle class.

Procedures. Each child was tested on the 125 items in two

15-20 minute individual testing sessions in a room other than his classroom. The child was told that the examiner (E) had a special book to show him and that E would like the child to help her read the book. The child was seated beside E at a small table. The book was laid flat on the table, and the following general instructions were given to the child:

> We are going to read this book together. I am going to read the words in the book. I would like you to help me read the book by putting your finger on the pictures I ask you to.

No warm-up items were used. E immediately began with the first item. When the book was open the stimulus pictures were located on the right of the 5 x 7 cards which made up the book pages, and the instructions for the examiner to read were on the left. For example,

> Here are two pictures. One is called the shoes are under the plane. One is called the shoes are over the plane. Put your finger on the picture called the shoes are under the plane.

The answers were scored as correct (1) or incorrect (0) by E. Each child's responses were recorded on a score sheet which contained spaces to check a response correct or incorrect. The raw data were then keypunched, and an item analysis was run.

Results

General summary data for the test are found in Table I. Table II provides individual item information.

TABLE I. Summary Date for the Preliminary Comparative Language Test

Number of Subjects	20
Number of Test Items	125
Test Mean	84.60
Test S.D.	17.63
Mean Item Difficulty	0.68
Mean Item-total Score Correlation	0.48
Standard Error of Correlation	0.23
Standard Error of Measurement	4.28
KR-20 Reliability	0.94

The items were grouped into seven categories, and the mean percent passing by category was calculated. This provided a rough hierarchy of category difficulty. Table III contains the results.

TABLE II. Item Analysis of the Preliminary Comparative Language Test

Item Number	Item Category	Item Content	Number Passing	Difficulty Index	Point Biserial	T-value
1	color	brown	19	0.950	0.168	0.724
2	color	white	15	0.750	0.558	2.850
3	position	beside	14	0.700	0.220	0.955
4	shape	rectangle	6	0.300	0.646	3.692
5	quantity	fewest	19	0.050	0.019	0.079
6	color	black	18	0.900	0.603	3.209
7	position	open	20	1.000	0.0	0.0
8	quantity	more	15	0.750	0.282	1.248
9	color	purple	15	0.750	0.611	3.278
10	shape	circle	17	0.850	0.463	2.215
11	size	little	20	1.000	0.0	0.0
12	relational	same	12	0.600	0.017	0.071
13	relational	big but less	7	0.350	0.414	1.927
14	shape	square	14	0.700	0.556	2.839
15	color and shape	orange circle	18	0.900	0.060	0.256
16	size	shorter	6	0.300	0.250	1.096
17	shape	triangle	16	0.800	0.076	0.322
18	color and shape	white triangle	14	0.700	0.518	2.570
19	quantity	fewer	8	0.400	0.399	1.847
20	size	biggest	19	0.950	-0.179	-0.771

TABLE II. Continued

Item Number	Item Category	Item Content	Number Passing	Difficulty Index	Point Biserial	T-value
21	color and shape	orange rectangle	16	0.800	0.476	2.294
22	size	narrower	5	0.250	0.611	3.278
23	position	closed	19	0.950	0.168	0.724
24	color and shape	yellow triangle	15	0.750	0.585	3.057
25	quantity	less	13	0.650	0.178	0.768
26	position	above	12	0.600	0.397	1.834
27	color	green	15	0.750	0.564	2.901
28	position	not in the corner	16	0.800	0.425	1.991
29	position	in the corner	17	0.850	0.064	0.270
30	quantity	less	13	0.650	0.379	1.740
31	position	out	19	0.950	0.662	3.749
32	quantity	many	2	0.100	0.367	1.672
33	size	tallest	17	0.850	0.447	2.117
34	shape	rectangle	10	0.500	0.466	2.232
35	size	taller	8	0.400	0.411	1.913
36	shape	circle	18	0.900	0.594	3.129
37	color	red	17	0.850	0.609	3.262
38	position	between	15	0.750	0.161	0.693
39	color and shape	purple triangle	14	0.700	0.245	1.072

TABLE II. Continued

Item Number	Item Category	Item Content	Number Passing	Difficulty Index	Point Biserial	T-value
40	size	wider	5	0.250	0.470	2.261
41	position	below	13	0.650	0.483	2.341
42	color	yellow	14	0.700	0.569	2.935
43	size	big	18	0.900	0.467	2.243
44	position	out	20	1.000	0.0	0.0
45	color and shape	green square	17	0.850	0.495	2.420
46	color and shape	purple square	14	0.700	0.626	3.406
47	position	over	12	0.600	-0.185	-0.800
48	size	bigger	5	0.250	0.618	3.336
49	color	blue	12	0.600	0.391	1.801
50	size	little	5	0.250	0.087	0.372
51	relational	less but big	8	0.400	0.441	2.083
52	relational	different	8	0.400	-0.100	-0.425
53	size	short	9	0.450	-0.131	-0.561
54	shape	triangle	11	0.550	0.453	2.154
55	size	littler	10	0.500	0.111	0.472
56	size	wide	7	0.350	0.011	0.047
57	color and shape	red circle	16	0.800	0.570	2.945
58	shape	square	15	0.750	0.571	2.952
59	color	orange	16	0.800	0.490	2.386
60	position	over	18	0.900	0.099	0.422

TABLE II. Continued

Item Number	Item Category	Item Content	Number Passing	Difficulty Index	Point Biserial	T-value
61	size	narrowest	10	0.500	0.343	1.551
62	color and shape	red rectangle	17	0.850	0.569	2.934
63	position	top	20	1.000	0.0	0.0
64	position	below	13	0.650	-0.011	-0.047
65	shape	circle	18	0.900	0.603	3.209
66	size	little	18	0.900	0.555	2.829
67	position	low	19	0.950	0.155	0.665
68	position	beside	14	0.700	0.385	1.768
69	position	closed	17	0.850	0.601	3.193
70	shape	circle	19	0.950	0.662	3.749
71	relational	more but shorter	8	0.400	0.067	0.283
72	quantity	most	11	0.500	0.529	0.642
73	color and shape	green triangle	15	0.750	0.087	0.372
74	position	under	18	0.900	0.458	2.184
75	position	up/down	15	0.750	-0.034	-0.143
76	position	up/down	11	0.550	0.225	0.978
77	position	under	18	0.900	0.177	0.761
78	position	top	19	0.950	0.662	3.748
79	relational	same in number	10	0.500	0.390	1.796
80	size	shortest	14	0.700	0.074	0.313

TABLE II. Continued

Item Number	Item Category	Item Content	Number Passing	Difficulty Index	Point Biserial	T-value
81	shape	square	12	0.600	0.551	2.802
82	position	in the corner	19	0.950	0.155	0.665
83	position	open	19	0.950	0.662	3.749
84	color and shape	black circle	15	0.750	0.484	2.345
85	position	between	16	0.800	0.490	2.386
86	shape	square	14	0.700	0.486	2.361
87	position	in	19	0.950	0.662	3.749
88	relational	different in number	11	0.550	0.061	0.259
89	color and shape	white square	17	0.850	0.593	3.126
90	position	not in the corner	18	0.900	0.080	0.339
91	relational	same	17	0.850	0.609	3.262
92	color and shape	brown rectangle	13	0.640	0.562	2.886
93	shape	triangle	15	0.750	0.302	1.346
94	position	bottom	17	0.850	0.455	2.166
95	relational	different and less	8	0.400	0.102	0.426
96	position	in	17	0.850	-0.108	-0.459
97	size	shorter	6	0.300	0.536	2.693
98	relational	little but less	12	0.600	-0.007	-0.030

TABLE II. Continued

Item Number	Item Category	Item Content	Number Passing	Difficulty Index	Point Biserial	T-value
99	relational	same in number	8	0.400	0.363	1.655
100	position	high	19	0.950	0.662	3.749
101	position	bottom	18	0.900	0.244	1.069
102	shape	rectangle	11	0.550	0.435	2.050
103	size	taller	5	0.250	0.484	2.345
104	position	over	18	0.900	0.050	0.214
105	shape	triangle	11	0.550	0.435	2.050
106	relational	different in number	8	0.400	0.233	1.016
107	color and shape	blue square	11	0.550	0.552	2.809
108	size	big	19	0.950	0.662	3.749
109	size	widest	13	0.650	0.373	1.707
110	size	wider	5	0.250	0.356	1.617
111	shape	rectangle	11	0.550	0.470	2.260
112	relational	different	17	0.850	0.544	2.753
113	position	below	16	0.800	-0.077	-0.328
114	position	low	17	0.850	0.601	3.193
115	quantity	fewer	5	0.250	0.403	1.869
116	relational	less but taller	5	0.250	0.329	1.479
117	color and shape	brown circle	16	0.800	0.636	3.494

TABLE II. Continued

Item Number	Item Category	Item Content	Number Passing	Difficulty Index	Point Biserial	T-value
118	quantity	less	8	0.400	0.072	0.308
119	quantity	more	10	0.500	0.361	1.641
120	color and shape	yellow triangle	16	0.800	0.512	2.529
121	position	high	19	0.950	0.662	3.749
122	quantity	more	14	0.700	0.080	0.340
123	size	tall	5	0.250	0.309	1.379
124	position	around/ through	13	0.650	0.428	2.010
125	position	forward/ backward	17	0.850	0.137	0.586

TABLE III. The Preliminary Comparative Language Test Mean Percent Passing by Content Category

N = 20

Category	Percent Passing
Position	86.3
Color	78.3
Color and Shape	76.3
Shape	68.1
Compound Relations	53.2
Size	50.3
Quantity	45.5

Discussion

While the overall procedures of the test and the test reliability proved satisfactory, a number of specific items were found to be in need of replacement or revision. All items with a negative point-biserial (a negative point-biserial results when performance on a given item correlates negatively with total test performance) were discarded or revised. The most evident reasons for the failure of these items to appropriately discriminate were: 1) perceptual confusion, i.e., some items were confusing because the differences between alternatives were too subtle, 2) the provision of additional cues, i.e., the child could determine the answer without taking the entire choice criterion into consideration, or 3) unclear instructions. Some items were changed in order to balance the two parallel forms which were subsequently built. Table IV summarizes the actions taken. The remaining 113 items were retained unchanged.

The pilot trial with the preliminary version of the test also confirmed that the test was too lengthy to be administered in a single session.

Although only a crude indication, the analysis of the content categories suggested that there was a hierarchy in the acquisition of the language concepts tested.

TABLE IV. Discarded or Revised Items*

Item	Concept	Point Biser.	T-value	Revised (R) Discarded (D)	Criterion
53	short	-.131	-.561	R	misleading perceptual cues
64	below	-.011	-.047	D	ambiguous
75	up/down	-.034	-.143	D	inconsist-ent with format
76	up/down	.225	.978	D	inconsist-ent with format
113	below	-.077	-.328	D	high scorers missed item
96	in	-.108	-.459	R	balance forms
47	over	-.185	-.800	D	ambiguous
52	different	-.100	-.425	D	ambiguous
98	little but less	-.007	-.030	R	misleading perceptual cues
118	less	.072	.308	R	poor wording
30	more	.379	1.740	R	balance forms
51	big but less	.441	2.083	R	balance forms and poor wording

*Item 20 had a negative point biserial, but after looking at the data it was decided to retain the item without revision because only one subject missed the item. Because that subject was a very high scorer on the test it unduly influenced the item results.

CHAPTER IV

THE MAIN STUDY

Purpose

The results of the preliminary study provided encouragement concerning the item format and overall procedures for measuring the comparative language of young children. It also suggested a number of ways in which measures could be improved. The second study was designed to 1) improve the construction of the Comparative Language Test, 2) investigate further the acquisition of the concepts tested, and 3) gain some evidence concerning the validity of the test. Additionally, since the preliminary study had indicated the length of the test to be excessive and since the test would be presumably useful as a pre- and posttreatment measure, two parallel forms were developed.

Method

Sample. A total of 52 children enrolled in four university early childhood education programs served as subjects for the study. The children ranged in age from two years, eight months, to five years, five months. There were 27 girls and 25 boys. Fourteen of the children had low socio-economic status (SES) backgrounds and were attending the school on scholarship. The remaining children were from middle-class families where the father was engaged in

professional, managerial, or student occupations.

Procedures. The revised version of the Comparative Language Test was administered to each of the children individually in two sessions in a room other than their classrooms. Thirty children were readministered the test 10-21 days after their initial testing sessions.

The children were told that the examiner (E) had a special book and that she would like the child to help her read it. E would read the words and the child would point to the picture when E asked him to. Other than a slight change in the introductory comments, the procedure from the preliminary study was retained.

The Comparative Language Test. The revised version of the Comparative Language Test consisted of 113 unchanged items from the preliminary version, six revised items from the preliminary version, and 43 new items. The resulting 162 items were arranged by categories into two parallel forms (A and B). When item information was available from the preliminary study, this information was used to develop comparability of the forms.

The item categories were retained from the original test. New items were added and some old items were revised so that approximately equal numbers of items in each category were provided on each test form.

The 81 5" x 7" cards of each form (A and B) were assembled

into book form and shown one at a time to the individual subjects. The remaining items, also in book form, were administered during the second session within the following week. When the book was opened the subject saw a stimulus picture on the right side. The left side contained the directions to be read by E.

Several sample items are included as Appendix A. A complete set of the stimulus pictures and their accompanying directions may be obtained from the author.

Results

A primary consideration of this study was to examine information relevant to test construction. The test was evaluated as a whole and in two parts (Forms A and B). Item analysis was run on individual items, and summary data were computed. This information is contained in Tables V and VI for the complete Comparative Language Test and in Tables VII, VIII, and IX for parallel forms A and B.

Although individual pilot items were discarded or revised on the basis of a negative point biserial and T-value, the criterion for item revision and/or rejection will now be based on a T-value of 2.021 ($p < .05$). These items have been marked with an asterisk in Tables VI, VIII, and IX. Table X breaks down poor items into categories and presents data relevant to that context.

TABLE V. Summary Data for the Complete Comparative Language Test

Number of Subjects	52
Number of Test Items	162
Test Mean	117.13
Test S.D.	25.49
Mean Item Difficulty	0.725
Mean Item-Total Score Correlation	0.593
Standard Error of Correlation	0.140
Standard Error of Measurement	4.38
KR-20 Reliability	0.970

TABLE VI. Item Analysis for Individual Items of the Complete Comparative Language Test

Item Number	Item Category	Item Content	Number Passing	Difficulty Index	Point Biserial	T-value
1	color	white	44	0.846	0.686	6.666
2	shape	square	39	0.750	0.653	6.094
3	quantity	more	50	0.962	0.472	3.783
4	color	brown	48	0.923	0.627	5.694
5	position	between	41	0.788	0.531	4.431
6	relational	same	41	0.788	0.498	4.038
7	position	not in the corner	44	0.846	0.203	1.466*
8	size	taller	21	0.404	0.499	3.555
9	position	above	43	0.827	0.567	4.864
10	color	white	44	0.846	0.686	6.666
11	size	little	52	1.000	0.0	0.0*
12	position	low	52	1.000	0.0	0.0*
13	relational	same and fewer	6	0.115	0.0	0.003*
14	color and shape	red circle	43	0.827	0.726	7.470
15	relational	different	44	0.846	0.282	2.082
16	size	big	52	1.000	0.0	0.0*
17	size	widest	30	0.577	0.391	3.002
18	position	below	42	0.808	0.116	0.822*
19	color and shape	blue square	41	0.788	0.699	6.914
20	shape	square	39	0.750	0.669	6.357

TABLE VI. Continued

Item Number	Item Category	Item Content	Number Passing	Difficulty Index	Point Biserial	T-value
21	quantity	few	32	0.615	0.491	3.986
22	color	brown	48	0.923	0.568	4.876
23	position	beside	40	0.769	0.651	6.065
24	position	in the corner	46	0.885	0.415	3.226
25	size	little	6	0.115	0.102	0.725*
26	color	black	48	0.923	0.452	3.579
27	position	closed	49	0.942	0.234	1.704
28	color and shape	purple triangle	37	0.712	0.551	4.671
29	size	wider	20	0.385	0.053	0.377*
30	position	high	49	0.942	0.360	2.732
31	position	bottom	49	0.942	0.422	3.291
32	shape	rectangle	32	0.615	0.476	3.823
33	size	bigger	17	0.327	0.387	2.969
34	color and shape	green rectangle	37	0.712	0.713	7.184
35	position	over	46	0.885	0.318	2.374
36	color	yellow	46	0.885	0.651	6.068
37	shape	circle	48	0.923	0.236	1.721
38	relational	fewer and shorter	19	0.365	0.378	2.889
30	quantity	fewer	22	0.423	0.040	0.281*
40	position	under	45	0.865	0.424	3.313

TABLE VI. Continued

Item Number	Item Category	Item Content	Number Passing	Difficulty Index	Point Biserial	T-value
41	position	top	44	0.846	0.351	2.654
42	relational	same in number	29	0.558	0.283	2.084
43	size	littler	20	0.385	0.571	4.919
44	color	red	47	0.904	0.462	3.687
45	relational	more and taller	40	0.769	0.150	1.071*
46	position	open	51	0.981	0.385	2.952
47	quantity	more	25	0.481	0.150	1.076*
48	color and shape	yellow square	41	0.788	0.671	6.406
49	shape	circle	45	0.865	0.316	2.355
50	size	tall	10	0.192	0.078	0.552*
51	color	black	50	0.962	0.342	2.576
52	color and shape	white rectangle	44	0.846	0.705	7.024
53	quantity	fewest	18	0.346	0.331	2.478
54	color	purple	45	0.865	0.632	5.766
55	size	big	9	0.173	0.067	0.477*
56	color and shape	brown circle	47	0.904	0.585	5.102
57	relational	more and different	25	0.481	0.543	4.572
58	size	narrow	23	0.442	-0.012	-0.087*

TABLE VI. Continued

Item Number	Item Category	Item Content	Number Passing	Difficulty Index	Point Biserial	T-value
59	color	yellow	45	0.865	0.689	6.730
60	size	shorter	20	0.385	0.391	3.006
61	shape	triangle	39	0.750	0.449	3.554
62	position	in	43	0.827	0.407	3.153
63	color	red	46	0.885	0.590	5.165
64	color and shape	black triangle	40	0.769	0.680	6.552
65	relational	fewer and bigger	9	0.173	0.402	3.108
66	color	purple	46	0.885	0.583	5.071
67	color	green	42	0.808	0.604	5.353
68	shape	rectangle	28	0.538	0.651	6.065
69	relational	different in number	27	0.519	0.123	0.875*
70	color	green	45	0.865	0.634	5.800
71	position	out	52	1.000	0.0	0.0*
72	color	blue	46	0.885	0.706	7.040
73	relational	little and fewer	30	0.577	0.194	1.397*
74	color and shape	orange circle	47	0.904	0.501	4.090
75	quantity	many	6	0.115	0.352	2.662
76	color	orange	41	0.788	0.586	5.119
77	size	short	13	0.250	-0.061	-0.429*

TABLE VI. Continued

Item Number	Item Category	Item Content	Number Passing	Difficulty Index	Point Biserial	T-value
78	color	blue	45	0.865	0.550	4.659
79	color	orange	43	0.827	0.627	5.684
80	size	narrower	15	0.288	0.396	3.052
81	shape	triangle	42	0.808	0.359	2.716
82	color	brown	47	0.904	0.583	5.068
83	position	high	51	0.981	0.155	1.106*
84	shape	rectangle	29	0.558	0.473	3.791
85	color	brown	44	0.846	0.632	5.760
86	size	narrowest	20	0.385	0.126	0.899*
87	relational	different and fewer	18	0.346	0.159	1.142*
88	color	purple	45	0.865	0.519	4.296
89	size	shorter	20	0.385	0.255	1.863
90	position	over	42	0.808	0.540	4.542
91	position	below	39	0.750	0.526	4.370
92	color and shape	orange rectangle	47	0.904	0.534	4.466
93	size	tallest	44	0.846	0.448	3.539
94	shape	square	42	0.808	0.650	6.041
95	color	purple	43	0.827	0.646	5.992
96	relational	same and more	13	0.250	0.000	0.003*
97	color	yellow	46	0.885	0.663	6.263

TABLE VI. Continued

Item Number	Item Category	Item Content	Number Passing	Difficulty Index	Point Biserial	T-value
98	relational	different	41	0.788	0.431	3.380
99	quantity	less	42	0.808	0.477	3.840
100	color and shape	red triangle	44	0.846	0.678	6.515
101	relational	more and bigger	46	0.885	0.389	2.987
102	shape	circle	46	0.885	0.694	6.811
103	position	top	49	0.942	0.344	2.593
104	relational	same	47	0.904	0.393	3.024
105	color	yellow	47	0.904	0.567	4.870
106	size	shortest	40	0.761	0.486	3.935
107	size	wide	8	0.154	0.023	0.162*
108	color	orange	43	0.827	0.613	5.480
109	position	low	46	0.885	0.295	2.181
110	position	open	49	0.942	0.108	0.769*
111	relational	less and taller	14	0.269	0.259	1.894
112	size	narrower	16	0.308	0.310	2.308
113	position	in	47	0.904	0.529	4.406
114	relational	more and little	35	0.673	0.330	2.473
115	shape	square	38	0.731	0.689	6.714
116	color	orange	43	0.827	0.567	4.864
117	color	white	43	0.827	0.589	5.149

TABLE VI. Continued

Item Number	Item Category	Item Content	Number Passing	Difficulty Index	Point Biserial	T-value
118	position	in the corner	44	0.846	0.425	3.316
119	size	littlest	51	0.981	0.325	2.428
120	color	white	46	0.885	0.576	4.978
121	color and shape	yellow rectangle	43	0.827	0.732	7.602
122	shape	circle	46	0.885	0.724	7.431
123	quantity	more	50	0.962	0.256	1.873
124	color	green	44	0.846	0.749	7.985
125	position	between	44	0.846	0.439	3.457
126	shape	triangle	43	0.827	0.028	0.200*
127	relational	different in number	28	0.538	0.54	4.543
128	position	above	32	0.615	0.217	1.569*
129	size	bigger	8	0.154	0.347	2.615
130	color	green	45	0.865	0.725	7.439
131	size	wider	19	0.365	0.336	2.522
132	shape	rectangle	34	0.654	0.578	5.006
133	color	red	43	0.827	0.509	4.180
134	position	not in the corner	41	0.788	0.319	2.377
135	quantity	most	34	0.654	0.059	0.420*
136	relational	more and shorter	29	0.558	0.304	2.256

TABLE VI. Continued

Item Number	Item Category	Item Content	Number Passing	Difficulty Index	Point Biserial	T-value
137	color and shape	blue triangle	46	0.885	0.330	2.473
138	size	biggest	51	0.981	0.166	1.187*
139	position	closed	50	0.962	0.134	0.959*
140	color	red	43	0.827	0.573	4.940
141	size	littler	13	0.250	0.467	3.738
142	color and shape	white square	49	0.942	0.600	5.301
143	color and shape	purple square	42	0.808	0.628	5.713
144	size	taller	21	0.404	0.405	3.129
145	quantity	fewer	15	0.288	0.546	4.610
146	quantity	more	27	0.519	0.041	0.292*
147	color	blue	44	0.846	0.627	5.697
148	color	blue	48	0.923	0.593	5.210
149	size	big	52	1.000	0.0	0.0*
150	position	out	52	1.000	0.0	0.0*
151	color and shape	green triangle	42	0.808	0.401	3.093
152	shape	triangle	36	0.692	0.543	4.571
153	color and shape	black circle	44	0.846	0.588	5.136
154	position	bottom	48	0.923	0.449	3.551
155	color	black	48	0.923	0.488	3.958

TABLE VI. Continued

Item Number	Item Category	Item Content	Number Passing	Difficulty Index	Point Biserial	T-value
156	size	little	52	1.000	0.0	0.0*
157	relational	same in number	30	0.577	0.527	4.382
158	color and shape	brown rectangle	38	0.731	0.607	5.400
159	quantity	few	14	0.269	0.124	0.886*
160	position	under	46	0.885	0.328	2.453
161	position	beside	36	0.692	0.464	3.708
162	color	black	46	0.885	0.309	2.296

* Items which need to be revised according to the 2.021 or less T-value criterion.

TABLE VII. Summary Data for Form A and Form B of the Comparative Language Test

	Form A	Form B
Number of Subjects	52	52
Number of Items	81	81
Test Mean	57.29	60.23
Test S.D.	13.04	12.99
Mean Item Difficulty	0.706	0.744
Mean Item-Total Score Correlation	0.594	0.612
Standard Error of Correlation	0.140	0.140
Standard Error of Measurement	3.14	3.06
KR-20 Reliability	0.942	0.944

TABLE VIII. Item Analysis for Individual Items--Form A

Item Number	Item Category	Item Content	Number Passing	Difficulty Index	Point Biserial	T-value
1	color	white	44	0.846	0.688	6.702
2	shape	square	39	0.750	0.697	6.879
3	quantity	more	50	0.962	0.472	3.788
4	color	brown	48	0.923	0.637	5.848
5	position	between	41	0.788	0.531	4.436
6	relational	same	41	0.788	0.488	3.954
7	position	not in the corner	44	0.846	0.210	1.517*
8	size	taller	21	0.404	0.496	4.036
9	position	above	43	0.827	0.587	5.127
10	color	white	44	0.846	0.688	6.702
11	size	little	52	1.000	0.0	0.0*
12	position	low	52	1.000	0.0	0.0*
13	relational	same and fewer	6	0.115	-0.059	-0.416*
14	color and shape	red circle	43	0.827	0.700	6.933
15	relational	different	44	0.846	0.271	1.991*
16	size	big	52	1.000	0.0	0.0
17	size	widest	30	0.577	0.392	3.014
18	position	below	42	0.808	0.168	1.205*
19	color and shape	blue square	41	0.788	0.737	7.716
20	shape	square	39	0.750	0.714	7.218

TABLE VIII.Continued

Item Number	Item Category	Item Content	Number Passing	Difficulty Index	Point Biserial	T-value
21	quantity	few	32	0.615	0.521	4.312
22	color	brown	48	0.923	0.582	5.060
23	position	beside	40	0.769	0.677	6.507
24	position	in the corner	46	0.885	0.405	3.132
25	size	little	6	0.115	0.094	0.664*
26	color	black	48	0.923	0.455	3.710
27	position	closed	49	0.942	0.208	1.503*
28	color and shape	purple triangle	37	0.712	0.567	4.876
29	size	wider	20	0.385	0.098	0.694*
30	position	high	49	0.942	0.347	2.616
31	position	bottom	49	0.942	0.398	3.064
32	shape	rectangle	32	0.615	0.496	4.044
33	size	bigger	17	0.327	0.431	3.377
34	color and shape	green rectangle	37	0.712	0.733	7.630
35	position	over	46	0.885	0.322	2.404
36	color	yellow	46	0.885	0.650	6.042
37	shape	circle	48	0.923	0.206	1.486*
38	relational	fewer and shorter	19	0.365	0.421	3.283
39	quantity	fewer	22	0.423	0.059	0.416*
40	position	under	45	0.865	0.445	3.515

TABLE VIII. Continued

Item Number	Item Category	Item Content	Number Passing	Difficulty Index	Point Biserial	T-value
41	position	top	44	0.846	0.332	2.492
42	relational	same in number	29	0.558	0.349	2.636
43	size	littler	20	0.385	0.571	4.913
44	color	red	47	0.904	0.472	3.790
45	relational	more and taller	40	0.769	0.159	1.140*
46	position	open	51	0.981	0.379	2.895
47	quantity	more	25	0.481	0.206	1.488*
48	color and shape	yellow square	41	0.788	0.694	6.815
49	shape	circle	45	0.865	0.290	2.139
50	size	tall	11	0.212	0.050	0.353*
51	color	black	49	0.942	0.341	2.562
52	color and shape	brown circle	43	0.827	0.721	7.327
53	quantity	fewest	18	0.346	0.365	2.774
54	color	purple	45	0.865	0.614	5.495
55	size	big	10	0.192	0.030	0.215*
56	color and shape	brown circle	48	0.923	0.560	4.777
57	relational	more and different	26	0.500	0.529	4.413
58	size	narrow	23	0.442	0.013	0.092*
59	color	yellow	45	0.865	0.678	6.530

TABLE VIII. Continued

Item Number	Item Category	Item Content	Number Passing	Difficulty Index	Point Biserial	T-value
60	size	shorter	19	0.365	0.436	3.430
61	shape	triangle	38	0.731	0.462	3.686
62	position	in	43	0.827	0.427	3.341
63	color	red	46	0.885	0.594	5.224
64	color and shape	black triangle	40	0.769	0.674	6.445
65	relational	fewer and bigger	10	0.192	0.420	3.268
66	color	purple	46	0.885	0.576	4.979
67	color	green	42	6.808	0.580	5.029
68	shape	rectangle	28	0.538	0.701	6.948
69	relational	different in number	26	0.500	0.202	1.459*
70	color	green	45	0.865	0.588	5.136
71	position	out	51	0.981	0.110	0.786*
72	color	blue	47	0.904	0.683	6.604
73	relational	little and fewer	31	0.596	0.193	1.387*
74	color and shape	orange circle	46	0.885	0.502	4.103
75	quantity	many	7	0.135	0.281	2.068
76	color	orange	40	0.769	0.614	5.503
77	size	short	13	0.250	-0.054	-0.380*
78	color	blue	45	0.865	0.530	4.431

TABLE VIII. Continued

Item Number	Item Category	Item Content	Number Passing	Difficulty Index	Point Biserial	T-value
79	color	orange	42	0.808	0.624	5.653
80	size	narrower	15	0.288	0.393	3.020
81	shape	triangle	41	0.788	0.376	2.871

*Items which need to be revised according to the 2.021 or less T-value criterion.

TABLE IX. Item Analysis for Individual Items--Form B

Item Number	Item Category	Item Content	Number Passing	Difficulty Index	Point Biserial	T-value
1	color	brown	47	0.904	0.578	5.014
2	position	high	51	0.981	0.164	1.177*
3	shape	rectangle	29	0.558	0.427	3.342
4	color	brown	44	0.846	0.636	5.820
5	size	narrowest	20	0.385	0.153	1.097*
6	relational	different and fewer	18	0.346	0.146	1.042*
7	color	purple	45	0.865	0.554	4.701
8	size	shorter	20	0.385	0.272	1.999*
9	position	over	42	0.808	0.557	4.745
10	position	below	39	0.750	0.575	4.963
11	color and shape	orange rectangle	47	0.904	0.503	4.116

TABLE IX. Continued

Item Number	Item Category	Item Content	Number Passing	Difficulty Index	Point Biserial	T-value
12	size	tallest	44	0.846	0.443	3.490
13	shape	square	42	0.808	0.602	5.336
14	color	purple	43	0.827	0.177	6.513
15	relational	same and more	13	0.250	0.024	0.169*
16	color	yellow	46	0.885	0.669	6.368
17	relational	different	41	0.788	0.470	3.762
18	quantity	less	42	0.808	0.493	4.011
19	color and shape	red triangle	44	0.846	0.660	6.215
20	relational	more and bigger	46	0.885	0.391	3.005
21	shape	circle	46	0.885	0.697	6.874
22	position	top	49	0.942	0.366	2.785
23	relational	same	47	0.904	0.433	3.394
24	color	yellow	47	0.904	0.578	5.014
25	size	shortest	40	0.761	0.455	3.611
26	size	wide	8	0.154	0.027	0.192*
27	color	orange	43	0.827	0.630	5.744
28	position	low	46	0.885	0.340	2.558
29	position	open	49	0.942	0.131	0.937*
30	relational	less and taller	14	0.269	0.263	1.930
31	size	narrower	16	0.308	0.303	2.251

TABLE IX. Continued

Item Number	Item Category	Item Content	Number Passing	Difficulty Index	Point Biserial	T-value
32	position	in	47	0.904	0.548	4.636
33	relational	more and little	35	0.673	0.372	2.836
34	shape	square	38	0.731	0.648	6.023
35	color	orange	43	0.827	0.584	5.081
36	color	white	43	0.827	0.595	5.238
37	position	in the corner	44	0.846	0.434	3.410
38	size	littlest	51	0.981	0.315	2.348
39	color	white	46	0.885	0.595	5.236
40	color and shape	yellow rectangle	43	0.827	0.740	7.782
41	shape	circle	46	0.885	0.739	7.751
42	quantity	more	50	0.962	0.265	1.946*
43	color	green	44	0.846	0.750	8.029
44	position	between	44	0.846	0.480	3.864
45	shape	triangle	43	0.827	0.036	0.251*
46	relational	different in number	28	0.538	0.581	5.045
47	position	above	32	0.615	0.218	1.579*
48	size	bigger	8	0.154	0.333	2.498
49	color	green	45	0.865	0.723	7.398
50	size	wider	19	0.365	0.322	2.403
51	shape	rectangle	34	0.654	0.539	4.524

TABLE IX. Continued

Item Number	Item Category	Item Content	Number Passing	Difficulty Index	Point Biserial	T-value
52	color	red	43	0.827	0.513	4.227
53	position	not in the corner	41	0.788	0.361	2.737
54	quantity	most	34	0.654	0.050	0.356*
55	relational	more and shorter	29	0.558	0.350	2.640
56	color and shape	blue triangle	46	0.885	0.345	2.597
57	size	biggest	51	0.981	0.175	1.257*
58	position	closed	50	0.962	0.158	1.128*
59	color	red	43	0.827	0.584	5.081
60	size	littler	13	0.250	0.445	3.510
61	color and shape	white square	49	0.942	0.608	5.411
62	color and shape	purple square	42	0.808	0.614	5.495
63	size	taller	21	0.404	0.408	3.159
64	quantity	fewer	15	0.288	0.541	4.549
65	quantity	more	27	0.519	0.076	0.542*
66	color	blue	44	0.846	0.623	5.635
67	color	blue	48	0.923	0.605	5.377
68	size	big	52	1.000	0.0	0.0*
69	position	out	52	1.000	0.0	0.0*
70	color and shape	green triangle	42	0.808	0.433	3.399

TABLE IX. Continued

Item Number	Item Category	Item Content	Number Passing	Difficulty Index	Point Biserial	T-value
71	shape	triangle	36	0.692	0.606	5.390
72	color and shape	black circle	44	0.846	0.599	5.284
73	position	bottom	48	0.923	0.466	3.728
74	color	black	48	0.923	0.489	3.960
75	size	little	52	1.000	0.0	0.0*
76	relational	same in number	30	0.577	0.531	4.428
77	color and shape	brown rectangle	38	0.731	0.598	5.281
78	quantity	few	14	0.269	0.119	0.850*
79	position	under	46	0.885	0.349	2.637
80	position	beside	36	0.692	0.448	3.545
81	color	black	46	0.885	0.322	2.402

*Items which need to be revised according to the 2.021 or less T-value criterion.

Reliability. Test-retest reliability was investigated for the complete Comparative Language Test. Of the 52 subjects tested, 24 were retested 10-21 days after the initial testing. A dependent t test was run on the means of the test ($\overline{X}$ = 125.42) and the retest ($\overline{X}$ = 126.13). There was no significant difference between the means ($t = -.701$, $df = 23$, $p > .05$). A Pearson product moment correlation

TABLE X. Unsatisfactory Items

Category	Unsatisfactory Items	Inappropriate Format	Too Easy	Other
Color	--	--	--	--
Shape	--	--	--	--
Color and shape	--	--	--	--
Position	8	--	7	1
Quantity	5	5	--	--
Size	13	9	4	--
Relational	6	--	--	6
Total	32	14	11	7

coefficient was also computed for test-retest. It yielded a correlation of .973, which is significantly different from zero at the .01 level.

Due to the length of the complete Comparative Language Test, two parallel forms were developed from the total 162 items. Each form (A and B) contained 81 items selected to approximate equivalency on the basis of the item analysis of pilot items. The means from Form A ($\bar{X}$ = 57.29) and Form B ($\bar{X}$ = 60.23) were compared by utilizing a dependent t test to determine whether the means were significantly different. The results indicated the mean from Form A was significantly different from the mean of Form B ($t = -4.5752$, $df = 50$, $p < .05$). A Pearson product moment correlation coefficient was computed for Form A and B. It produced a correlation of .9365 ($p < .01$).

Pearson product moment correlation coefficients were also computed correlating the total test with each of the forms. Form A correlated with the total test at .9736 ($p < .01$). Form B correlated with the total test at .9828 ($p < .01$).

An additional measure of reliability available was the Kuder Richardson 20 internal consistency reliability. The KR-20 reliability for the total Comparative Language Test was .970. Form A had a KR-20 of .942 and Form B had a KR-20 of .944.

Validity. The essential question of test validity is how well the test measures what it is built to measure. In the problem statement three hypotheses were formulated concerning validity estimates. The first stated that within the sample, there would be a significant relation between the overall score on the Comparative Language Test and chronological age. In order to test this hypothesis, the 52 subjects were divided into two age groups. Group I ($n=17$) contained subjects ranging in age from two years, eight months, to four years, three months (mean age = three years, nine months). Group II ($n=35$) contained subjects ranging in age from four years, five months to five years, five months (mean age = four years, 11 months). Mean test scores for both groups were determined ($\bar{X}$ Group I = 98.47, $\bar{X}$ Group II = 126.20). A two-tailed Behrens Fisher t' test was run on the means, and it was determined that the test means for the two groups were significantly different ($t' = 3.769$, $df' = 24$, $p < .05$).

The second hypothesis stated that within the sample there would be a significant correlation between the overall score on the Comparative Language Test with I.Q. or mental age. Twenty-nine of the 52 subjects were administered the Peabody Picture Vocabulary Test within two months of testing with the Comparative Language Test. Computation of a Pearson product moment correlation coefficient yielded a correlation coefficient of .427 ($p < .05$).

The third hypothesis stated that middle-class children would score significantly higher than lower SES children. The 52 subjects were divided into two SES groups, low SES (n=14) and middle-class (n=38). The mean test score for each group was determined ($\bar{X}$ low SES = 89.50, MC = 127.32). A two-tailed Behrens Fisher t' test was run on the data, and it was determined that the two means were significantly different ($t' = 5.2844$, $df' = 17$, $p < .05$).

Sex differences were also analyzed. Computations specified a mean score for males (n=25) of 112.16 and a mean score for females (n=27) of 121.74. A Behrens Fisher t' test indicated no significant differences existed ($t' = 1.356$, $df' = 40$, $p > .05$).

Hierarchy Determination. An additional area of interest concerned the possible existence of a hierarchy of difficulty among content categories. Pilot data seemed to indicate a hierarchical trend (see Table III). The mean percent passing was computed for each category of items. An analysis of variance for unequal n's (n = items per category) was run and followed up by a Newman-Keuls

test to determine significant differences between means. The results appear in Table XI.

TABLE XI. Analysis of Variance for Difficulty of Item Categories

Category	Mean Difficulty	Source	df	MS	F
Position	87.3	Between	6	.6909	21.0*
Color	86.9	Within	155	.0329	
Color and Shape	82.5				
Shape	75.2				
Compound Relational	54.9				
Quantity	53.7				
Size	49.6				

* $p < .001$

The Newman-Keuls analysis indicated that no significant differences existed among the mean item difficulties for position, color or color and shape, nor were there significant differences among the last four categories in Table XI. However, the two groupings of categories (categories 1, 2, 3, vs. 4, 5, 6, 7) differ significantly from each other in mean item difficulty.

Discussion

The purpose of this study was to: 1) build a test that would meet standard requirements of reliability and validity, 2) be practical as an assessment tool for children ages three to five, and 3) be helpful in defining a hierarchy of difficulty of acquiring comparative language.

The test seems to have very adequately met the standard psychometric requirements of reliability and validity. Reliability measures were high and significant. The formulation of two parallel forms of the test was successful and extremely important for further administration and revision of the test. Each form correlated highly and significantly with the total test and with its alternate form. Form A seems to be slightly more difficult, and in later administrations an adjusted form might be utilized.

Hypotheses concerning validity were supported. As expected, those children who were older did significantly better on the test. Since the acquisition of language is an evolving process, older children should have more comparative language. Socioeconomic differences were also expected since the literature indicated that the context in which language is learned is related to the type of language learned. Since middle-class children learn more elaborated language codes (Bernstein, 1968), they were expected to perform better on the Comparative Language Test. The results supported this hypothesis.

The correlation between the Comparative Language Test and the Peabody Picture Vocabulary indicated that, although the tests were positively correlated, the correlation coefficient was small enough to suggest that the two tests were not measuring the same content. They have 19 percent of the variance in common.

The test, in its two parallel forms, seems to be a practical measurement tool for children ages three through five. Each form (81 items) takes approximately 15 minutes to administer. The test is in book form, which is familiar to most children. The pictures are colorful, and directions are short. The individual children being tested are asked to participate in the "activity" but are not required to verbalize.

The administration of the test is not difficult since the examiner is simply required to read the short text opposite each stimulus picture. The scoring is not difficult as it merely requires a check in an appropriate column labeled right or wrong beside the item number. Each column may then be totaled for the child's raw score.

One major problem seems to exist as the result of the particular format of some of the items. The following item provides an example:

> Here are three leaves. One is called wide. Another is called wider. Another is called the widest. Put your finger on the leaf called "wider." (Item #29)

In this type of format three relative concept variables are presented. The problem is that there is actually more than one

reasonable and correct answer. The middle-sized leaf (keyed the correct choice in this item) is wider than the narrowest leaf, but the widest leaf is also "wider" than the middle-sized leaf. Wider, then, is a term that is appropriate in the comparison of two objects only. Likewise, "wide" is typically considered as an opposite to "narrow," and all three leaves may realistically be considered "wide." The three-choice alternative presented with the item stimulus picture seems most appropriate when considering the "est" variable word.

The evolution of a hierarchy of difficulty of comparative language is suggested by the data. With the exception of the categories of quantity and size, the data of this study replicate that of the preliminary study. In both cases the order of difficulty was found to be: position, color, color and shape, shape, compound relational, quantity and size (or size and quantity). However, two points should be noted. First, the problem noted in size items may have erroneously increased the difficulty of these items. Secondly, the analysis in the main study indicated that the mean item difficulty between adjacent categories is not always significant. Rather, the data provide only a crude division between less difficult items (position, color, color and shape) and more difficult items (shape, compound relational, quantity and size).

CHAPTER V

SUMMARY

The purpose of the study was to begin the development of a test of comparative language that would meet standard requirements of reliability, validity, and practicality. Additionally, it was hoped to begin the definition of a hierarchy of difficulty in the acquisition of comparative language.

Characteristics of the Test

Reliability. The test seems to have very adequately met the psychometric requirements of reliability. All reliability coefficients (internal consistency and test-retest) exceeded .90.

Validity. Preliminary checks on the validity of the test were encouraging. Expected differences between age and SES groupings were confirmed. A significant correlation between the total Comparative Language Test score and the Peabody Picture Vocabulary Test (IQ) score was found as expected. However, these two tests share only 19 percent common variance, which suggests that they are measuring somewhat different aspects of cognitive development.

Practicality. With few exceptions, the format and procedures of the test were found acceptable and convenient to use. Two

parallel forms were developed, each comparing very favorably with the total test. The parallel forms may be administered in approximately 15 minutes by a relatively inexperienced tester.

The item analysis suggested a number of additional revisions which should be undertaken (see Table X). Additionally, a logical next step would be the detailed analysis of the category subscales. Enlargement of these scales may be necessary to provide reliable diagnostic information.

Further investigation of the validity of the test is essential. A number of suggestive areas has been described in Chapter II. Comparative language of the sort measured here ought to predict performance on visual discrimination tasks, conservation tasks, and a variety of other problem-solving tasks. The relationship to mathematics and reading readiness should be explored. Additionally, much needs to be learned concerning the relation of comparative language to other aspects of language development.

Hierarchy of Comparative Language Development

The investigation of a possible hierarchy of difficulty of comparative language tends to support the trend found in the preliminary study with the exception of the quantity and size categories which were reversed. The analysis of category difficulty indicated that although the hierarchial trend held up, the mean item difficulty between adjacent categories is not always significantly

different. Instead, a crude division between less difficult and more difficult categories was found.

Additional investigation of a possible comparative language hierarchy is necessary. Poor items would need to be revised prior to this since those items tended to distort the hierarchial analysis. If a developmental hierarchy of comparative language can be ascertained, it would prove extremely valuable in test development, curriculum construction, and in the study of language acquisition.

REFERENCES

Bailey, B. L. A crucial problem in language intervention as it relates to the disadvantaged. IRCD Bulletin Supplement, Summer, 1966, 11, 3A.

Baratz, J. C. Language in the economically disadvantaged child: a perspective. Journal of the American Speach and Hearing Association, 1968, 10, 143-145.

Beilin, H. Learning and operational convergence in logical thought development. Journal of Experimental Child Development, 1965, 2 (4), 317-338.

Berko, J. and Brown, R. Psycholinguistic research methods. In P.H. Mussen (Ed.), Handbook of Research Methods in Child Development. New York: Wiley, 1960, 517-557.

Bernstein, B. A. A socio-linguistic approach to social learning. In J. L. Frost (Ed.), Early Childhood Education Rediscovered. New York: Holt, Rinehart and Winston, Inc., 1968, 445-466.

Blank, M. Some philosophical influences underlying preschool intervention for disadvantaged children. In F. Williams (Ed.), Language and Poverty. Chicago: Markham Publishing Company, 1970, 62-81

Braine, M.D.S. The ontogeny of certain logical operations: Piaget's formulation examined by nonverbal methods. Psychology Monographs, 1959, 73 (5).

Braine, M.D.S. Development of a grasp of transitivity of length: a reply to Smedsluna. Child Development, 1964, 35 (3), 799-810.

Bussis, A. Children's understanding of basic language structure. In Cognitive Growth in Preschool Children. Educational Testing Service Memorandum, 1968.

Distar. Language Program. Chicago: Science Research Associates, 1969.

Engelmann, S. How to construct effective language programs for the poverty child. In F. Williams (Ed.), Language and Poverty. Chicago: Markham Publishing Company, 1970, 102-122.

Evans, E. D. Contemporary Influences in Early Childhood Education. New York: Holt, Rinehart and Winston, Inc., 1971.

Frost, J. L. Developing literacy in young children. In J. L. Frost (Ed.), Early Childhood Education Rediscovered. New York: Holt, Rinehart and Winston, Inc., 1968, 371-374.

Frost, J. L. and Rowland, T. Cognitive development and literacy in disadvantaged children: a structure-process approach. In J. L. Frost (Ed.), Early Childhood Education Rediscovered. New York: Holt, Rinehart and Winston, Inc., 1968, 374-401.

Griffiths, J. A., Shantz, C. A. and Sigel, I. E. A methodological problem in conservation studies: the use of relational terms. Child Development, 1967, 38, 841-848.

Green, G. E. Note on conservation: methodological and definitional considerations. Child Development, 1966, 37, 977-983.

Halasa, O. A developmental study of number conservation attainment among disadvantaged children. Paper delivered at the American Educational Research Association annual meeting, Los Angeles, 1969.

Hess, R. D. and Shipman, V. Early experience and the socialization of congnitive modes in children. Child Development, 1965, 36, 869-886.

Inhelder, B., Bovet, M., Sinclair, H., and Smock, C. D. On cognitive development. American Psychologist, 1964, 21, 160-164.

Kamii, C. An application of Piaget's theory to the conceptualization of a preschool curriculum. Paper presented at a conference entitled "Conceptualizations of Preschool Curricula," sponsored by the Department of Educational Psychology, The City University of New York, May, 1970.

Lavatelli, C. S. Piaget's Theory Applied to an Early Childhood Curriculum. Boston: American Science and Engineering, Inc., 1970.

Menyuk, P. Language theories and educational practices. In F. Williams (Ed.), Language and Poverty, Chicago: Markham Publishing Company, 1970, 190-212.

Montemerlo, M. D. The effects of training and ability on conservation of number. Unpublished doctoral dissertation, The Pennsylvania State University, 1970.

Nimnicht, G. P. The New Nursery School. New York: General Learning Corporation: Early Childhood Division, 1969.

Palmer, F. H. Minimal intervention and intellective change at two and three. Unpublished research report, State University of New York at Stony Brook, 1968.

Peters, D. L. Piaget's conservation of number: the interaction of language comprehension and analytic style with three methods of training. Unpublished doctoral dissertation, Stanford University, 1968.

Piaget, J. Language and Thought of the Child. New York: Meridian Books, The World Publishing Company, 1926.

Sinclair-de-Zwart, H. Developmental psycholinguistics. In D. Elkind and J. H. Flavell(Eds.), Studies in Cognitive Development: Essays in Honor of Jean Piaget. New York: Oxford University Press, 1969, 315-336.

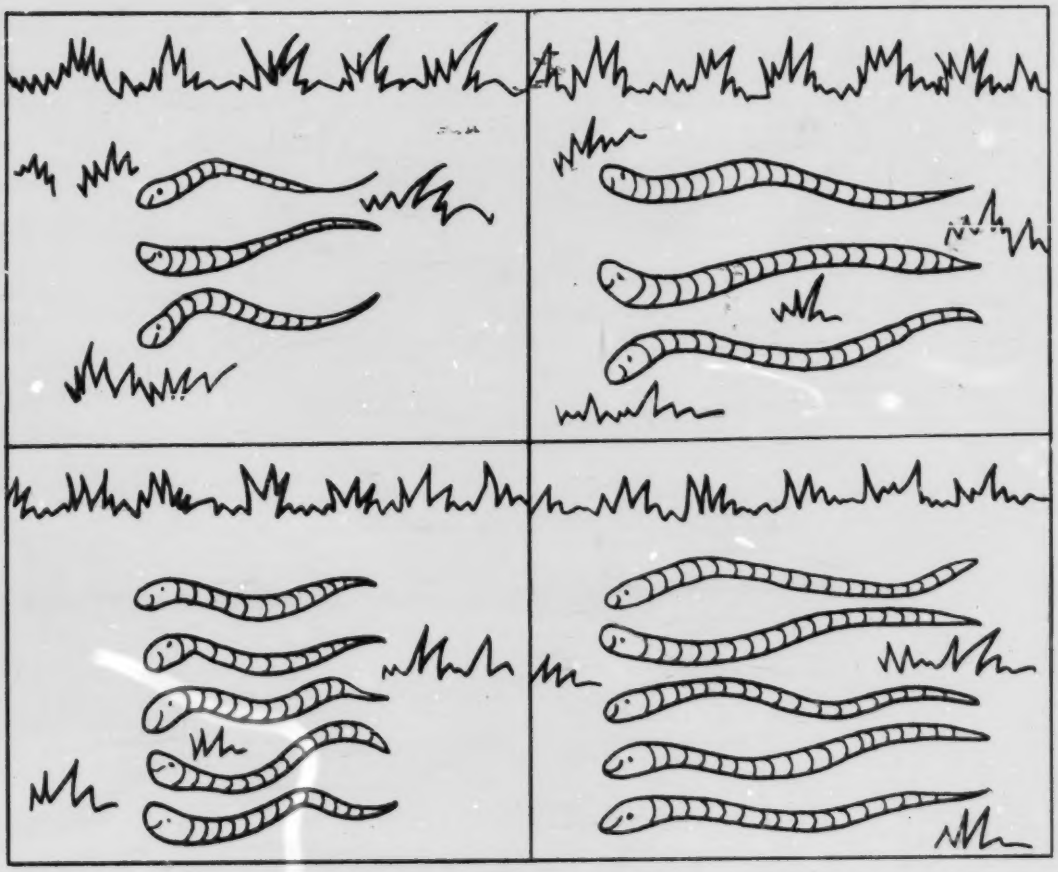
Williams, F. Some preliminaries and prospects. In F. Williams (Ed.), Language and Poverty. Chicago: Markham Publishing Company, 1970, 1-11.

Wohlwill, J. F. A study of the development of number concept by scalogram analysis. Journal of Genetic Psychology, 1960, 97, 356-377.

APPENDIX

APPENDIX A

SAMPLE TEST ITEMS

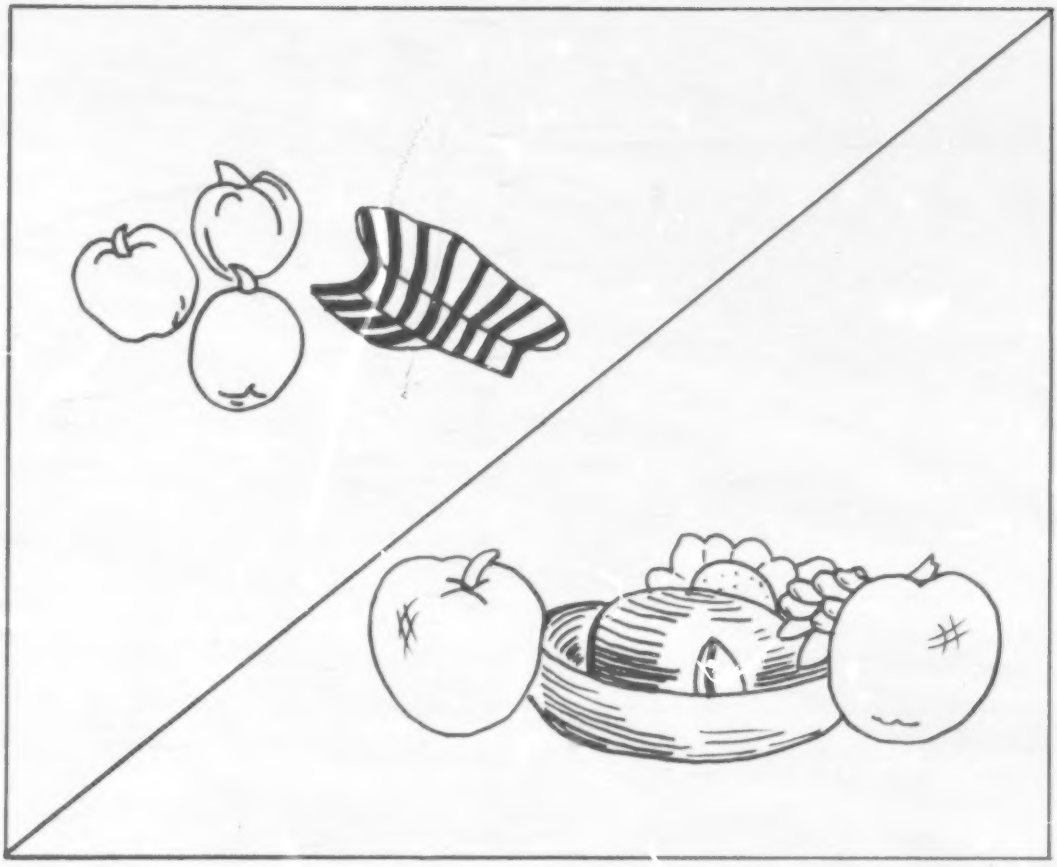


Here are four pictures.

Look carefully at each of them (point).

Put your finger on the picture which

shows more and shorter worms.



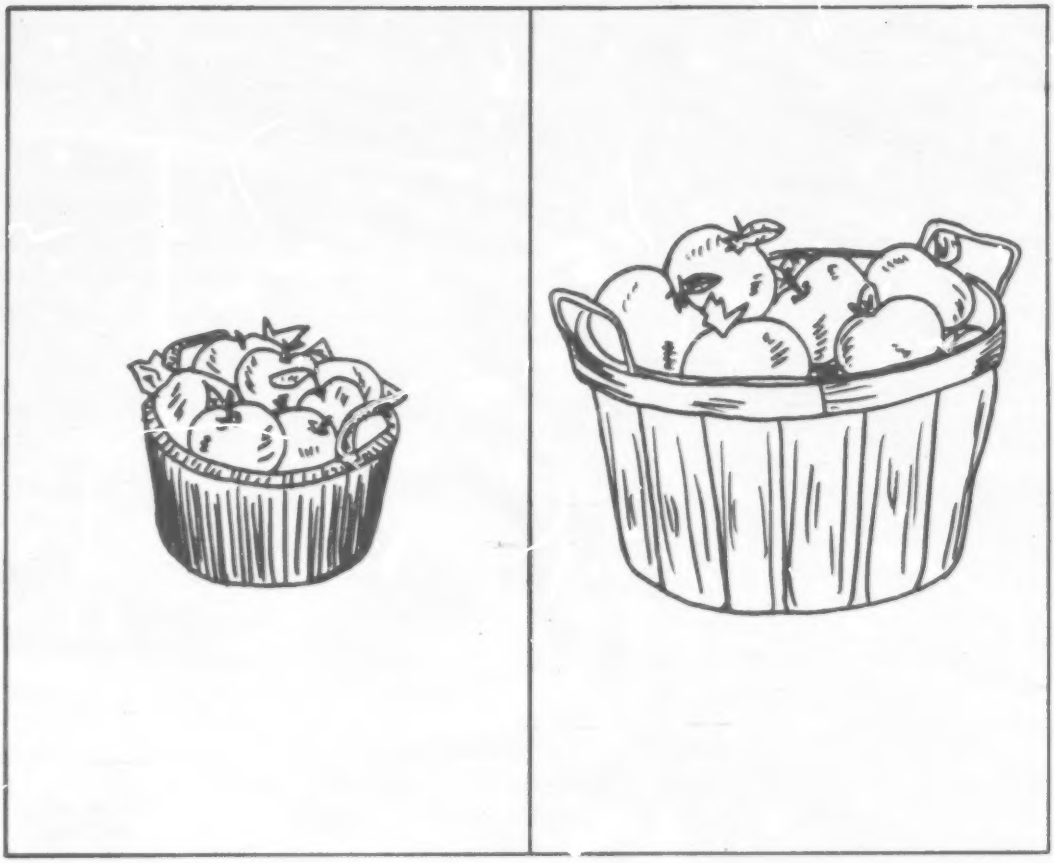
Here are two pictures.

One shows a hat beside some apples.

One shows a hat between some apples.

Put your finger on the picture that

shows a hat between some apples.



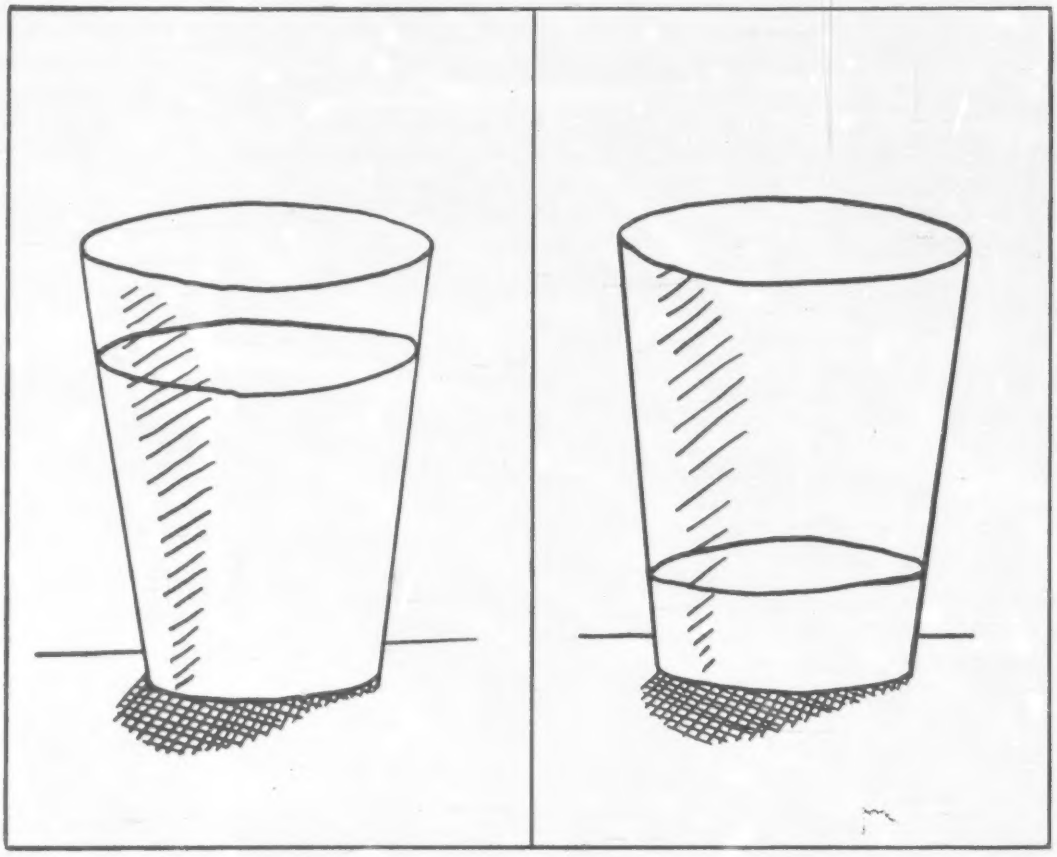
Here are two pictures.

One is called little basket of apples.

One is called big basket of apples.

Put your finger on the picture called

little basket of apples.



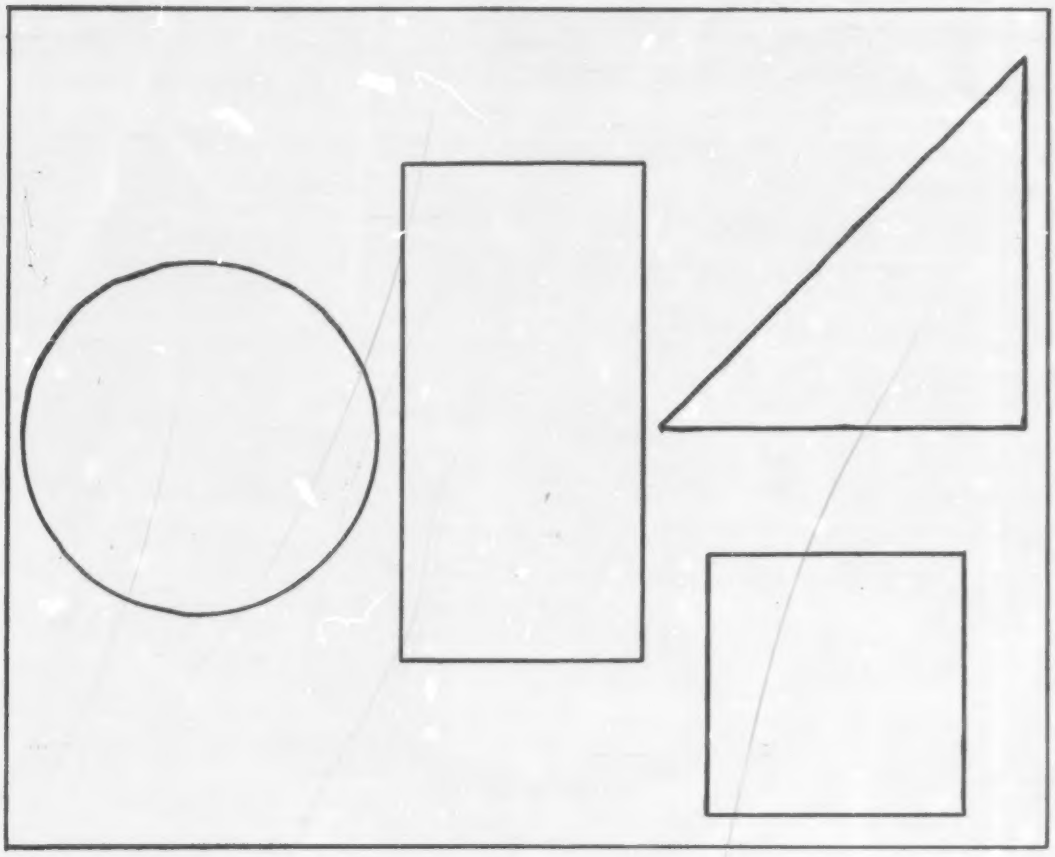
Here are two pictures.

One is called more milk

One is called less milk.

Put your finger on the picture

called less milk.



Here are some shapes.

Put your finger on the rectangle.